DIGITALCOMMONS @WAYNESTATE

Journal of Modern Applied Statistical Methods

Volume 9 | Issue 1 Article 11

5-1-2010

Beyond Alpha: Lower Bounds for the Reliability of Tests

Nol Bendermacher

Radboud University, Nijmegen, The Netherlands, Bendermacher@hotmail.com

Journal of Modern Applied Statistical Methods
May 2010, Vol. 9, No. 1, 95-102

1538 – 9472/10/$95.00

Beyond Alpha: Lower Bounds for the Reliability of Tests

Nol Bendermacher
Radboud University,
Nijmegen, The Netherlands

The most common lower bound to the reliability of a test is Cronbach's alpha. However, several lower bounds exist that are definitely better, that is, higher than alpha. An overview is given as well as an algorithm to find the best: the greatest lower bound.

Key words: test reliability; Cronbach's alpha.

Introduction

The concept of reliability is based on the notion of accuracy or precision of a measurement. This article is confined to the reliability of tests - psychological or other - consisting of a number of items and to the situation where a test is administered only once. A person's score on such a test is the sum of his/her scores on the individual items.

According to classical test theory, the score x_{ij} of person i on item j consists of two parts: the true score τ_{ij} and an error component ε_{ij}: $x_{ij} = \tau_{ij} + \varepsilon_{ij}$. Moreover, classical test theory assumes that the error components are uncorrelated with the true parts as well as with each other. As a consequence the covariance matrix Γ of the items is the sum of two components: the covariance matrix (Γ_τ) of the true parts and the covariance matrix (Γ_ε) of the error components:

$$\Gamma = \Gamma_\tau + \Gamma_\varepsilon$$

The assumption of uncorrelated errors implies that Γ_ε is a diagonal matrix; thus the off-diagonals of Γ and Γ_τ are identical. The assumption of independent errors is essential for all measures discussed herein. Many conditions exist which lead to the violation of the assumption of independent errors, for example: in a test with a time limit where an unanswered item results in a minimum score, the errors of the last items may correlate, or in a long or difficult test errors may become correlated due to the effect of fatigue or declining motivation during the test administration.

The reliability of a test consisting of v items is defined as:

$$\rho_{tt} = 1 - \frac{\sigma_e^2}{\sigma_t^2} \qquad (1)$$

where σ_e^2 is the error variance and σ_t^2 is the total variance of the test scores:

$$\sigma_t^2 = \sum_{i=1}^{v}\sum_{j=1}^{v}\Gamma_{ij} \qquad (2)$$

$$\sigma_e^2 = TR(\Gamma_e) = \sum_{i=1}^{v}\Gamma_{eii} \qquad (3)$$

Based on these formulae the definition of reliability can be rewritten as:

$$\rho_{tt} = 1 - \frac{\sigma_e^2}{\sigma_t^2} = 1 - \frac{\sum_{i=1}^{v}\Gamma_{eii}}{\sum_{i=1}^{v}\sum_{j=1}^{v}\Gamma_{ij}} \qquad (4)$$

Nol Bendermacher is a retired member of the Research Technical Support Group of the Faculty of the Social Sciences. Email him at: Bendermacher@hotmail.com.

It should be noted that this definition leaves undecided whether the unique variances (item variance components not correlated with any other item) are treated as error or as true variance. The lower bounds discussed herein are lower bounds according to both definitions.

Lower Bounds

If no other assumptions are added to those of the classical model it is impossible to assess the reliability of a test from a single administration; only lower bounds can be derived. From (4) it is clear that - given covariances Γ - the reliability is maximal if the trace of the error covariance matrix Γ_e is minimal. As Jackson and Agunwamba (1977) remarked, the only restrictions that the classical model imposes on the elements of Γ_ε are

(1) $0 \leq \Gamma_{eii} \leq \Gamma_{ii}$, and

(2) $\Gamma_\tau = \Gamma - \Gamma_e$ is non-negative definite.

Thus, if the set of values Γ_e that maximizes its trace $\sum_{i=1}^{v} \Gamma_{eii}$ under these restrictions can be located, the result would give the smallest possible value for the reliability given the covariance matrix Γ; this value is the greatest possible lower bound to the reliability. Jackson and Agunwamba (1977) and ten Berge, Snijders and Zegers (1981) described algorithms to find this largest lower bound; however, several well-known lower bounds are first put forth.

Guttman (1945) introduced a series of lower bounds called λ_1 through λ_6.

$$\lambda_1: \ 1 - \frac{\sum_{i=1}^{v} \Gamma_{ii}}{\sum_{i=1}^{v}\sum_{j=1}^{v} \Gamma_{ij}} = \frac{\sum_{i\neq}^{v}\sum_{j}^{v} \Gamma_{ij}}{\sum_{i=1}^{v}\sum_{j=1}^{v} \Gamma_{ij}} \qquad (5)$$

This λ_1 is the sum of the off-diagonal cells in Γ divided by the sum of all cells. The larger the item covariances, as compared to the variances, the larger λ_1.

$$\lambda_2 = \lambda_1 + \frac{\sqrt{\frac{v}{v-1}\sum_{i\neq}^{v}\sum_{j}^{v} \Gamma_{ij}^2}}{\sum_{i=1}^{v}\sum_{j=1}^{v} \Gamma_{ij}} \qquad (6)$$

Because $\lambda_2 \geq \lambda_1$, λ_2 should always be preferred over λ_1.

$$\lambda_3: \ \frac{v}{v-1}\lambda_1 = \frac{v}{v-1}\left(1 - \frac{\sum_{i=1}^{v} \Gamma_{ii}}{\sum_{i=1}^{v}\sum_{j=1}^{v} \Gamma_{ij}}\right) \qquad (7)$$

This λ_3 is better known as Cronbach's alpha. Guttman (1945) remarked "λ_3 is easier to compute than λ_2, since only the total variance and the item covariances are required. If the covariances are all positive and homogeneous, then λ_3 will not be much less than λ_2 and may be an adequate lower bound. If the covariances are heterogeneous, and in particular, if some are negative, then λ_2 will be definitely superior to λ_3. λ_2 can be positive and useful when λ_3 is negative and useless" (pp. 274-275). In brief, $\lambda_1 \leq \lambda_3 \leq \lambda_2$. Therefore, with modern computational facilities, λ_2 should always be preferred over λ_3. In actual practice, however, researchers tend to use λ_3, which is better known as Cronbach's alpha or, with dichotomous items, the Kuder-Richardson 20 (KR20).

Ten Berge and Zegers (1978) showed that λ_3 and λ_2 are members of a series of bounds μ_0, μ_1, μ_2, ..., defined by the following general formula:

$$\mu_r = \frac{1}{\sigma_t^2}\left(p_0 + \left(p_1 + \left(p_2 + \ldots\left(p_{r-1} + \left(p_r\right)^{\frac{1}{2}}\right)^{\frac{1}{2}} \ldots\right)^{\frac{1}{2}}\right)^{\frac{1}{2}}\right), \qquad (8)$$

where $r = 0, 1, 2, \ldots$.

$$p_h = \sum_{i\neq j} \sigma_{ij}^{(2^h)}, \ h = 1, 2, \ldots, r\text{-}1$$

$$p_h = \frac{v}{v-1}\sum_{i\neq j}\sigma_{ij}^{(2^h)},\ h = r$$

From this formula it is observed that $\mu_0 = \lambda_3 =$ Cronbach's alpha = KR20 and $\mu_1 = \lambda_2$. The differences between μ_{r+1} and μ_r rapidly converge to zero, thus, there is not much use in going further than μ_3.

$$\lambda_4\text{: } 2\left(1-\frac{\sigma_1^2+\sigma_2^2}{\sigma_t^2}\right) \qquad (9)$$

where σ_1^2 and σ_2^2 are the variances of two test halves:

$$\sigma_1^2 = \sum_i\sum_j\Gamma_{ij}$$

where i and j run over the items in the first test half and similarly,

$$\sigma_2^2 = \sum_i\sum_j\Gamma_{ij}$$

with i and j running over the items in the second test half.

A problem with λ_4 is that many ways exist by which to split a test into two parts, meaning that there are many different values for λ_4: the most interesting of them is the largest. In the statistical package SPSS (release 15.0.0) the value of λ_4 depends on the order of the items in the scale: it assigns the first v/2 items (with odd v the first (v+1)/2) to the first test half and the remaining items to the second half.

A simple algorithm to find a good split is based on the following: Imagine that the rows and columns of the covariance matrix are rearranged such that the items of the first test half come first, σ_1^2 and σ_2^2 are the sums of the upper left and the lower right quarter of the covariance matrix Γ respectively. Because the sum (σ_t^2) of the entire matrix is fixed, λ_4 is maximal if the sum of the lower left (and the upper right) quarter is maximal. This leads to the following algorithm:

1) Locate the pair of items with the highest covariance and assign one of them to test 1 and the other to test 2.
2) Try each ordered pair (i, j) of items not yet assigned. Compute the covariance between the two test parts if item i is assigned to test 1 and item j to test 2. After all pairs are tried, make the assignment that resulted in the highest covariance between the tests.
3) Repeat step 2 until all items have been assigned to one of the test-halves. In the case of an odd number of items, the last item is added to the group for which the mean covariance with the item is the smallest.

Given a specific split, Jackson and Agunwamba (1977) described a method to determine whether the resulting value of λ_4 is the greatest possible lower bound. Define: b = a vector with v-elements, with $b_i = 1$ if item i belongs to test half 1 and $b_i = -1$ if it belongs to test half 2; $A = \Gamma b$; $\theta_i = b_i A_i$, i = 1, v (this θ is the vector with error variances); and $\Gamma_t = \Gamma - \text{diag}(\theta_i)$. If Γ_t is non-negative definite and all $\theta_i \geq 0$, λ_4 is the greatest possible lower bound.

$$\lambda_5\text{: } \lambda_1 + \frac{2\sqrt{\max_i(\Gamma_{*i})}}{\sum_{i=1}^{v}\sum_{j=1}^{v}\Gamma_{ij}} \text{ with } \Gamma_{*i} = \sum_{j=1}^{v}\Gamma_{ij}^2 - \Gamma_{ii}^2 \qquad (10)$$

As Guttman (1945) noted, this measure will be larger than λ_2 if one item has large covariances with the other items compared with the covariances among those items. Otherwise λ_5 is less than or equal to λ_2.

$$\lambda_6\text{: } 1-\frac{\sum_{i=1}^{v}\Gamma_{ii}\left(1-P_i^2\right)}{\sum_{i=1}^{v}\sum_{j=1}^{v}\Gamma_{ij}} = 1-\frac{\sum_{i=1}^{v}\frac{1}{\Gamma_{ii}^{-1}}}{\sum_{i=1}^{v}\sum_{j=1}^{v}\Gamma_{ij}} \qquad (11)$$

where Γ_{ii}^{-1} denotes the i^{th} diagonal of the inverse of Γ. In these formulae P_i^2 is the squared multiple correlation in the multiple regression of

item i on the remaining v-1 items: $P_i^2 = 1 - \frac{1}{P_{ii}^{-1}}$. ($P_{ii}^{-1}$ denotes the i^{th} diagonal of the inverse of the correlation matrix from Γ).

Guttman (1945) explained that λ_6 will be larger than λ_2 if the multiple correlations are relatively large as compared to the zero-order correlations. Otherwise λ_6 will tend to be less than or equal to λ_2. Jackson and Agunwamba (1977) reported that λ_6 should be particularly advantageous in the fairly typical situation where the inter-item correlations are positive, moderate in size and somewhat similar. Jackson and Agunwamba (1977) added a seventh bound, called λ_7:

$$\lambda_7: 1 - \sum_{i=1}^{v} \frac{\sigma_i^2}{\sigma_t^2} + \frac{\sqrt{\frac{v}{v-1}\sum_{i\neq}\sum_{j} d_{ij}^2}}{\sigma_t^2} \qquad (12)$$

where d_{ij}^2 is defined as follows:

g = the value of j for which $\frac{\Gamma_{ij}^2}{\Gamma_{jj}}$ is largest

k = the value of i≠j for which $\frac{\Gamma_{ij}^2}{\Gamma_{ii}}$ is largest

r_{ij} = the correlation between items i and j

$$d_{ij}^2 = \sigma_i^2 \sigma_j^2 \max\left(r_{ig}^2 r_{kj}^2, r_{ij}^2\right)$$

Jackson and Agunwamba remarked that this bound will be substantially better than λ_2 when there is considerable variation among the squared correlations.

Woodhouse and Jackson (1977) showed some partial orders in the bounds λ_1 through λ_7: $\lambda_1 \le \lambda_3 \le \lambda_2 \le \lambda_7$, $\lambda_1 \le \lambda_4$, $\lambda_1 \le \lambda_5$, $\lambda_1 \le \lambda_6$. Table 1 shows a covariance matrix of four items and the lower bounds discussed for the reliability of their sum.

Table 1: Variances and Covariances of Four Items and Lower Bounds for the Reliability of their Sum

5.6 0.2 6.7 2.8 3.9 8.8 -1.2 1.9 3.0 10.8	λ_1 = 0.3992 $\lambda_2 = \mu_1$ = 0.5867 $\lambda_3 = \alpha = \mu_0$ = 0.5323 λ_4 = 0.5574 λ_5 = 0.6125 λ_6 = 0.5817 λ_7 = 0.5904 μ_2 = 0.5936 μ_3 = 0.5957

When Experimental Independence Does Not Hold

Guttman (1953) provided some lower bounds for the situation where the assumption of independent errors does not hold by introducing an additional quantity δ, for which in some specific situations upper bounds can be defined. Such situations are tests with a time limit and more general tests where the completion of an item depends on the completion of its predecessor. The adjusted measures are:

$$\lambda_1^* = \lambda_1 - \frac{\delta}{\sigma_t^2} \qquad (13)$$

$$\lambda_2^* = \lambda_2 - \frac{\delta}{\sigma_t^2} \qquad (14)$$

$$\lambda_3^* = \frac{v}{v-1}\lambda_1^* \qquad (15)$$

$$\lambda_4^* = \lambda_4 - \frac{2\delta}{\sigma_t^2} \qquad (16)$$

For the situation where the assumption of uncorrelated errors is violated only by the fact that the completion of an item depends on the completion of its predecessor, Guttman (1955) gives three upper bounds for δ, assuming that an item that is omitted results in the lowest possible score. Thus, the assumption of uncorrelated errors is weakened to the following: "The basic assumption from now on is that, if person i attempts item j, then his score on any later item

g (g > j) will be experimentally independent of his score on this attempted item j. That is, we are considering here the case where dependence is due solely to *omissions,* so that if a part is attempted, no further experimental dependence holds" (Guttman, 1955, p. 119). Defining

v' = the number of items with a non-zero variance

l_i = the minimum score on item i, also the score for an unattempted item

h_i = the maximum score on item i

$m_i = h_i - l_i$

x_i = mean score on item i with l_i subtracted

p_i = proportion of persons that attempt item i

the (estimates of the) upper bounds for δ are:

$$d_1 = \sum_{i=1}^{v'} \sqrt{m_i x_i (1-p_i) \sum_{j=i+1}^{v'} m_j} \qquad (17)$$

If $m_i = 1$ for all i, this formula reduces to:

$$d_1 = \sum_{i=1}^{v'} (v'-i) \sqrt{x_i (1-p_i)} \qquad (18)$$

$$d_2 = 2 \sum_{i=1}^{v'} \left(m_i \sqrt{1-p_i} \sum_{j=i+1}^{v'} \sqrt{m_j x_j} \right) \qquad (19)$$

If $m_i = 1$ for all i, this formula reduces to:

$$d_2 = 2 \sum_{i=1}^{v'} \left(\sqrt{1-p_i} \sum_{j=i+1}^{v'} \sqrt{x_j} \right) \qquad (20)$$

$$d_3 = 2 \sum_{i=1}^{v'} \sum_{j=i+1}^{v'} e_{ij} \qquad (21)$$

where

$$e_{ij} = \max\left(m_i \sqrt{m_j x_j (1-p_i)}, \; m_j \sqrt{m_i x_i (1-p_j)} \right) \qquad (22)$$

From these formulas it is clear that d_3 is at least as high as d_1 and d_2.

Finding the Greatest Lower Bound

Woodhouse and Jackson (1977) described an algorithm that finds the greatest lower bound (GLB) for the reliability of a test if only the assumptions of classical test theory hold. However, ten Berge, Snijders and Zegers (1981) showed that this algorithm will not always produce the correct lower bound. They described another algorithm that avoids these shortcomings and also is less time consuming. The algorithm, as implemented in this study, proceeds as follows: Define:

C = the given covariance matrix

C_0 = C-Diag(C)

and

R_i = The i^{th} row of C

1) Construct a v by r matrix T with r <= v and not too small. Ten Berge, Snijders and Zegers (1981) advised that r = the number of non-negative eigenvalues in Γ_0. In order to be safe, choose r = v.

   Similar to Bentler & Woodward (1980) the cells T_{ij} of T are defined as follows:

   If i > j T_{ij} is set to $-\frac{1}{\sqrt{i}}$

   If i = j T_{ij} is set to $\frac{1}{\sqrt{i}}$

   If i < j T_{ij} is set to 0

   (By this choice all rows have length 1.)

2) Perform the following steps for each row i of T:

   2.1) Compute $a = \text{MIN}(0, \sqrt{R_i^T T T^T R_i})$; a is the provisional estimate of the true variance of item i.

2.2) If $0 < C_{ii} < a$, row i is replaced by $\frac{-1}{C_{ii}} T^T R_i$

If $0 < a \leq C_{ii}$, row i is replaced by $\frac{-1}{a} T^T R_i$.

2.3) If a = 0, rescale row i to length 1.

3) Compute the (estimated) sum of error variances: $E = TR(T^TCT)$ and check for convergence. The process has converged if the following conditions hold:

a) E has not (sufficiently) decreased since the last check
b) All rows of T have length ≥ 1

If the process has not converged go back to step 2.

4) Compute the resulting estimate of Γ_t by copying C and replacing C_{ii} by $\text{MIN}\left(C_{ii}, \sqrt{R_i^T TT^T R_i}\right)$, i = 1,v and check whether its smallest eigenvalue is zero. If not, the whole procedure should be repeated with another starting value of T, but we wonder if such a situation will ever occur.

5) Define $E = C_{ii} - C_{tii}$ and estimate

$$\text{GLB} = 1 - \frac{E}{\sum_{i=1}^{v}\sum_{j=1}^{v} C_{ij}} \qquad (23)$$

If this algorithm is applied to the example of Table 1 the result is GLB = 0.7324. Ten Berge and Sočan (2004) provide several sources from which other programs can be obtained that compute the greatest lower bound.

The Effect of Sampling Error

A problem exists with several of the lower bounds described in the preceding text. When estimated from a small sample, λ_4, λ_5, λ_6, λ_7 and GLB will capitalize on chance; meaning that their estimates from the sample tend to overestimate the true population values. As Shapiro and ten Berge (2000) remarked: "It is well known that the g.l.b., based on small samples (even a sample of one thousand subjects is not generally enough) may severely overestimate the population value, and statistical treatment of the bias has been badly missing" (p. 413). They show that bias tends to increase with decreasing sample size and with lower values of GLB. Moreover, the bias is expected to be larger with more parameters to be estimated, that is, with greater v.

In absence of an analytical solution the use of brute (computing) force is suggested. The following bootstrapping approach could be used:

1) Compute from the sample covariance matrix C the selected lower bound, G_0. If the sample from which C is computed is available, steps 2 through 5 may be skipped and the sample plays the role of X in step 6.

2) Generate a n by v matrix F, filled with drawings from a standard normal distribution; n must be not too small and always larger than v: 1,000 or 2,000 is adequate.

3) Rotate the columns of F to orthogonality and scale them to mean 0 and length $\sqrt{n}$; F will act as the set of components from a principal components analysis.

4) Perform a principal components analysis on C, resulting in a diagonal matrix Λ with eigenvalues and the matrix V with the corresponding eigenvectors. Compute the factor matrix $A = V\Lambda^{\frac{1}{2}}$ and make sure that A is square; add zero columns if needed.

5) Construct the matrix $X = FA^T$. The resulting X has a multivariate normal distribution with covariance matrix $\frac{1}{n} X^T X = C$.

6) Draw k random samples from X. For each estimate the covariance matrix and the

chosen lower bound. The sampling consists of random selections (with replacement) of rows from X. Compute the mean G_m and the standard deviation s_m of these lower bound estimates. The standard error $\frac{s_m}{\sqrt{k}}$ may be used as a stop criterion during the simulations.

7) The difference G_0-G_m is an estimator of the bias by capitalization on chance and G_0 is corrected by taking $2G_0$-G_m instead. The correction may not be perfect, but it will be close if the sample is not too small and G_0 not too great.

In the example of Table 1 and assuming a sample size n = 200 the bias is estimated as 0.002839; taking n = 100 the bias estimate becomes 0.002942. These bias estimates are very small, possibly due to the small number of items.

A computer program, called Reliab, that computes some of the lower bounds to the reliability, including the GLB, is available at http://www.ru.nl/socialewetenschappen/rtog /software/statistische/kunst/

The Factor Analytic Approach

Factor analysis explains the correlations between a set of items by a limited set of underlying latent variables, called factors. The model allows the estimate for scores of individuals on the factors as weighted sums of their item scores. In this model it is possible not only to find lower bounds, but also to find real estimates of the reliability of the estimated factor scores from a single test administration.

In factor analytic models, the variance of an item is viewed as composed of two parts:

1) Common variance, i.e. variance that is shared with other items, and

2) Unique variance (or unicity), i.e. variance that is unique for the item: it consists of specificity and genuine error.

Defining:

Z = n × v matrix of standardized scores (z-scores) of n individuals on v items.

F = n × v matrix of true scores of the individuals on f factors; F is unknown.

B_z = v × f matrix of weights to estimate the factor scores F from the item scores Z: $\hat{F} = ZB$

Assume that the weights are scaled such that the variances of $\hat{F}$ (i.e., diagonal values of $\frac{1}{n}\hat{F}^T\hat{F}$) are unity. Thus if

A = v × f factor pattern, i.e., the matrix containing the weights of the factors in the reconstruction of Z: $Z = FA^T$ + error + unicities;

S = v × f factor structure; it contains the correlations between Z and F;

U = v × v diagonal matrix with unicities; and

R_{ff} = f × f matrix with correlations between the factors;

then the correlations between the factors and the factor score estimates are:

$$\begin{aligned} R_{f\hat{f}} &= \frac{1}{n}F^T\hat{F} \\ &= \frac{1}{n}F^T ZB_z \\ &= \frac{1}{n}F^T\left(FA^T + U\right)B_z \\ &= \frac{1}{n}F^T FA^T B_z \\ &= R_{ff}A^T B_z \\ &= S^T B_z \end{aligned} \qquad (24)$$

If this model is adhered to, the latent factors play the role of true scores, and although they are latent, $R_{f\hat{f}}$ contains estimates of the correlations between them and the factor score

estimates. The squares of these correlations can be interpreted as the reliabilities of the factor score estimates. This measure is also called the "factor determinacy" (McDonald, 1974, p. 213).

In this context two remarks must be made:

1) Factors, as they result from a factor analysis, are not completely defined: they function as axes in an f-dimensional space and any other set of f axes in that space will explain the correlations between the items equally well. Therefore, the orientation of the factors must be selected on the basis of additional criteria, for example their interpretability from a given theory, and

2) As with all regression models, the squared correlations between factors and factor score estimates tend to be inflated, especially when the analysis is based on a small sample.

Conclusion

A number of lower bounds to the reliability of a test have been discussed; all are based on the covariance matrix of the items in the test. It is clear that the most commonly used measure, known as Cronbach's alpha, KR20 or λ_3, is a poor choice; its only advantage over Guttman's λ_2 is its ease of computation by hand.

It is clear that - under the assumptions of the classical test theory and without additional assumptions - the measure known as the Greatest Lower Bound is the highest possible lower bound. Its only weakness, one shared with several of the other measures, is its sensibility to capitalization on chance if it is estimated from a relatively small sample. In the absence of analytical methods to correct this bias a bootstrapping approach using brute computing force is suggested in order to minimize the bias.

References

Bentler, P. M., & Woodward, J. A. (1980). Inequalities among lower bounds to reliability: with applications to test construction and factor analysis. *Psychometrika*, *45*, 249-267.

Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, *10*, 255-282.

Guttman, L. (1953). Reliability formulas that do not assume experimental independence. *Psychometrika*, *18*, 225-239.

Guttman, L. (1955). Reliability formulas for noncompleted or speeded tests. *Psychometrika*, *20*, 113-124.

Jackson, P. H., & Agunwamba, C. C. (1977). Lower bounds for the reliability of the total score on a test composed of non-homogeneous items I: Algebraic lower bounds. *Psychometrika*, *42*, 567-578.

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, *2*, 151-160.

McDonald, R. P. (1974). The measurement of factor indeterminacy. *Psychometrika*, *39*, 203-222.

Shapiro, A., & ten Berge, J. M. F. (2000). The asymptotic bias of minimum trace factor analysis with applications to the greatest lower bound to reliability. *Psychometrika*, *65*, 413-425.

Ten Berge, J. M. F., & Zegers, F. E. (1978). A series of lower bounds to the reliability of a test. *Psychometrika*, *43*, 575-479.

Ten Berge, J. M. F., Snijders, T. A. B., & Zegers, F. E. (1981). Computational aspects of the greatest lower bound to the reliability and constrained minimum trace factor analysis. *Psychometrika*, *46*, 201-213.

Ten Berge, J. M. F., & Sočan. (2004). The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. *Psychometrika*, *69*, 613-625.

Woodhouse, B., & Jackson, P. H. (1977). Lower bounds for the reliability of the total score on a test composed of non-homogeneous items: II: A search procedure to locate the greatest lower bound. *Psychometrika*, *42*, 579-591.